

Asymptotically Optimal Resource Block Allocation With Limited Feedback

Ilai Bisritz, *Student Member, IEEE*, Amir Leshem, *Senior Member, IEEE*

Abstract

Consider a channel allocation problem over a frequency-selective channel. There are K channels (frequency bands) and N users such that $K = bN$ for some positive integer b . We want to allocate b channels (or resource blocks) to each user. Due to the nature of the frequency-selective channel, each user considers some channels to be better than others. The optimal solution to this resource allocation problem can be computed using the Hungarian algorithm. However, this requires knowledge of the numerical value of all the channel gains, which makes this approach impractical for large networks. We suggest a suboptimal approach, that only requires knowing what the M -best channels of each user are. We find the minimal value of M such that there exists an allocation where all the b channels each user gets are among his M -best. This leads to feedback of significantly less than one bit per user per channel. For a large class of fading distributions, including Rayleigh, Rician, m-Nakagami and others, this suboptimal approach leads to both an asymptotically (in K) optimal sum-rate and an asymptotically optimal minimal rate. Our non-opportunistic approach achieves (asymptotically) full multiuser diversity and optimal fairness, by contrast to all other limited feedback algorithms.

Index Terms

Resource Allocation, Multiuser Diversity, Channel State Information, Random Bipartite Graphs.

I. INTRODUCTION

The problem of multiple users accessing a common medium is a central issue in every communication network [2]–[4]. The most common access schemes in practice are orthogo-

The authors are with the Faculty of Engineering, Bar-Ilan University, Ramat-Gan, Israel, e-mail: ilaibist@gmail.com, leshema@biu.ac.il.

This research was supported by the Israel Ministry of Science and Technology under grant 3-13038 and by a joint ISF-NRF research grant number 2277/16.

Parts of this paper were presented at the 2018 Wireless Communications and Networking Conference [1].

nal transmission techniques such as TDMA, FDMA , CDMA and OFDMA [5], [6]. Modern communication networks allocate each user a number of resource blocks, each consisting of several OFDMA subcarriers that need not be adjacent [7].

When using a frequency division technique in a frequency-selective channel, not all channels have the same quality. Furthermore, since different users have different positions, the channel gains of users in a given frequency band are independent. This introduces a resource allocation problem between users and frequency bands (channels). Theoretically, if the exact knowledge of all these channel gains were available, an optimal solution for this resource allocation could be computed. The performance gain that can be achieved by solving this problem over a frequency allocation that ignores the channel gains is known as multiuser diversity [8].

In FDD systems, the base station can only have knowledge of the channel gain of each user if all users transmit this information directly. In TDD systems, the base station can estimate the channel gain of a given user from the reciprocity of the channel. However, the base station must estimate the channel gains in all frequencies, not just the currently used ones, so a wideband transmission is required from each user. These transmissions need to be coordinated between users so that they will not collide. Moreover, although the channels are indeed reciprocal, the RF circuits are not, which calls for a calibration process [9].

As networks grow and bandwidth expands, the number of channels increases significantly. This is typical of OFDMA systems that tend to have a large number of subcarriers. For example, in LTE [7] there are from 128 to 2048 subcarriers in the downlink for each user. This implies that reporting the channel state information (CSI) to the base station (BS) inflicts a large communication overhead that makes such a scheme infeasible.

This paper deals with a channel allocation task with K channels where N users simultaneously demand b channels for each. This occurs when users are backlogged; i.e., always have packets to transmit. This represents a non-opportunistic approach where the base station waits for a queue of N users to form and only then computes a good allocation for all the users in the queue. When users leave the network and others replace them, the allocation is updated accordingly.

We propose a novel limited feedback scheme that significantly reduces feedback overhead compared to state of the art techniques [10]–[14]. The proposed technique provides a provably both asymptotically optimal sum-rate **and** max-min fairness (which none of the current methods

provide). The proof applies to a wide range of fading distributions that includes Rayleigh, Rician and m-Nakagami. Our technique uses ~ 0.85 bits per channel per user as of for $K = N = 20$, and ~ 0.52 bits per user per channel for $K = 512$ and $N = 128$.

In our approach, users only need to feed back the indices of their M -best channels, and not their numerical values. If M is large enough, there is a high probability of an allocation of users to channels where each user gets b channels, all of which are from his M -best channels. We prove that for $M \geq (b + 1)(1 + \varepsilon) \ln K$, for $\varepsilon > 0$, this probability approaches one as $K \rightarrow \infty$. While our approach is asymptotic, the probability that such a perfect matching allocation exists exceeds 99% for $K = 20$ channels and $N = 20$ users.

Figure 1 presents a toy example of our system with $N = K = 4$. Each user transmits the indices of his M best channels to the base station. The base station constructs a preferences matrix, computes a perfect matching in the resulting bipartite graph and transmits back the allocated channel to each user.

The existence of a perfect matching allocation relies on the theory of random bipartite graphs. We define a novel random bipartite graph where each user is represented by b vertices (“agents”) in the graph, all of which are connected to the M -best channels of this user on the other side of the graph. We prove that this bipartite graph has a perfect matching, with high probability. Due to its dependent edges, our random bipartite graph is not an Erdős–Rényi graph. Consequently, proving the existence of a perfect matching requires revisiting the original proof. Our proof is in the spirit of the original result by Erdős and Rényi for random bipartite graphs [15]. A generalization of our results to the case of unequal numbers of channels between users can be achieved using the same techniques. It is omitted due to the complicated formalism. However, we present detailed simulations of the CDF of the rates (not just the means) that indicate that our results are also valid for an unequal number of channels between users.

Throughout this paper, we use the term “channel”. This may represent several united subcarriers in the case where the coherence bandwidth is large and they are highly correlated (e.g., a resource block). In other words, a “channel” is a designed unit that should optimally be the smallest independent resource block in terms of the statistics of the corresponding channel gain.

A. Related Work

There is extensive literature on OFDMA resource allocation (see the surveys in [16], [17] and the references therein). These allocations deal with the subcarrier assignment and the power allocation over these subcarriers. The state of the art algorithms achieve a close to optimal sum-rate with reduced computational complexity compared to the optimal solution. In many studies, it has been shown that an equal power allocation between the subcarriers of each user achieves very close to optimal performance [18]–[23]. Hence, the problem of subcarrier allocation appears to be much more significant than that of power allocation. However, all these approaches assume that users estimate their CSI perfectly and feed all of it back to the BS. While this assumption may be reasonable for small networks, it is highly infeasible for large networks with a large number of subcarriers, such as LTE [7].

In order to overcome the need for the whole CSI to exploit the selectivity of the channel, many works have suggested suboptimal schemes with limited feedback. For an excellent overview of limited feedback systems, see [24]. The state of the art algorithms consider an opportunistic scheduling approach, where each available channel is assigned to a user with a high instantaneous channel gain [10]–[12], [25].

In [10]–[12] each user transmits one bit per channel to indicate if the corresponding channel gain exceeds a threshold, while in [25] each user transmits the highest-rate modulation scheme he can support out of a discrete set of schemes (which requires more than one bit per channel). In [26], [27], channels are grouped together into clusters and only the information about the quality of each cluster is transmitted by the user. This grouping approach limits the amount of multiuser diversity that can be achieved, so that the reduced feedback has a performance cost. This cost is avoided if the grouped subcarriers are highly correlated, but in this case the multiuser diversity is smaller. We argue that for highly correlated channels, using grouping cannot be considered a real feedback reduction, since assigning two bits to two highly correlated subcarriers is a waste of feedback to begin with. In this work we use the term “channel” to refer to the smallest group of uncorrelated subchannels, which can be considered a kind of grouping.

In [13], the grouping and thresholding approaches were combined to achieve a feedback of less than one bit per user per channel, which leads to a compromise of suboptimal performance (even

asymptotically). In [14], the M -best channels approach was used for opportunistic scheduling. Users are arranged in a queue, and each of them transmits the indices his M -best channels to the BS. The BS sequentially assigns each user the best channel from these M channels if it is available, and a random channel if none of them are. The thresholding approach requires the BS or the users to know the fading distribution to compute the optimal threshold. The M -best channels approach does not require any such knowledge, which makes it more robust.

All these approaches are opportunistic in nature and as such suffer from an inherent unfairness. Some of the methods have suggested obtaining some degree of fairness by designing the scheduler [10]–[12], [27], at the cost of a reduced sum-rate. In order to get a comparable sum-rate with a proportional fairness scheduler, each user needs to wait until he is chosen as the “opportunistic user”. Therefore, fairness can only be maintained by introducing very large delays for users, which is far from being desirable or practical.

The rationale behind the opportunistic approach is to exploit the multiuser diversity. By contrast, here we argue that this widespread approach is too conservative and that achieving multiuser diversity should not come at the expense of fairness. We argue that fairness is undermined because all these approaches schedule a user for each channel separately. Somewhat surprisingly, when analyzing all the channels together, each user is very likely to be “an opportunistic choice” for some channel. This calls for a more sophisticated mathematical analysis of the random matchings between users and channels as we employ in this paper, in contrast to the simplistic arguments dealing with one user at a time.

In contrast to [11]–[14], our work is not limited to the case of Rayleigh-fading channels, but applies to a much broader class of exponentially-dominated tail distributions. Besides Rayleigh fading, this class includes Rician, m -Nakagami and others.

In [28]–[30], a distributed channel allocation for Ad-Hoc networks was introduced, based on a carrier-sensing multiple access (CSMA) scheme and a thresholding approach for the channel gains. Although considering a different scenario, the analysis in [28]–[30] also used a random bipartite graph, albeit a very simplified one. The random bipartite graph in [28]–[30] was an Erdős–Rényi graph (i.e., existence of edges is independent), for which perfect matching existence results are well known [15]. There are no perfect matching existence results that can be employed for our graph, so a novel analysis is required. Our random bipartite graph differs from an

Erdős–Rényi graph in two key ways. First, we use the M -best channels approach instead of the thresholding approach. This means that each user node is connected to exactly M channel nodes, and not only in average. Second, we allow for an allocation of $b > 1$ channels for each user, which leads to b identical vertices. These two characteristics mean that the edges in our random bipartite graph are not independent. Moreover, the analysis in [28]–[30] was limited to bounded channel gains and only proved the asymptotic optimality of the expected sum-rate. In contrast, we prove asymptotic optimality for the random sum-rate in probability. More importantly, we prove asymptotic optimality for the random minimal rate in probability.

The existence of perfect matching allocations is closely related to the pure Nash equilibrium (NE) of interference games. In our previous work [31] we designed a game where all pure NE are an almost (or an exact) perfect matching between users and channels. We assumed that each user was allocated a single channel. The results of this paper can be exploited to generalize our previous results to the case of $b > 1$ channels per user.

B. Outline

The remainder of this paper is organized as follows. In Section II we formulate our channel allocation problem and present our limited feedback scheme. In Section III we prove that as $K \rightarrow \infty$, the probability that a perfect matching exists in the user-channel graph approaches one for $M \geq (b + 1)(1 + \varepsilon) \ln K$ (Theorem 2) and that it cannot approach one for smaller M (Lemma 3). ***This perfect matching is an allocation where each user gets b of his M -best channels.*** In Section IV we analyze the feedback requirements of our scheme and propose an efficient encoding method. In Section V we prove that a perfect matching allocation asymptotically attains ***both an optimal sum-rate and max-min fairness*** (Theorem 8). Section VI provides simulation results that support our analytical findings, compares our algorithm to state of the art algorithms and shows that our results hold for a more general model. Finally, Section VII concludes the paper.

II. PROBLEM FORMULATION

In our model, there are N users that are served by a BS that has K channels for allocation. Figure 1 depicts a toy example of our system with $K = N = 4$. We assume that $K = bN$

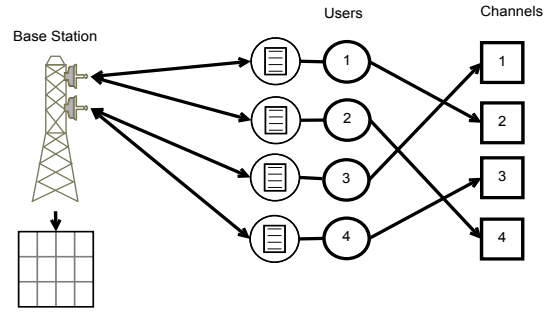


Fig. 1. The System

for some positive integer b and that each user is allocated b channels. The parameter b , which we treat as given, can be optimized beforehand. A generalization of our analysis to a different number of channels per user is possible using a more cumbersome proof. We later demonstrate in simulations that our scheme works just as well in this case. Throughout the paper, we use calligraphic capitals to denote sets (and events), bold capitals to denote matrices and capitals or lower case letters to denote scalars. Our assumptions on the channel gains are as follows:

Channel Gains Assumptions

- 1) For each user n , the channel gains $g_{n,1}, \dots, g_{n,K}$ are K i.i.d random variables. Channel gains of different users are independent. This is a widespread assumption in the channel allocation literature [12]–[14]. However, in contrast to [11]–[14], we do not assume Rayleigh fading or that the channel gains of different users are identically distributed.
- 2) Each user can perfectly measure each of his K channel gains. This widespread assumption [11]–[14], [26], [27] is necessary so that each user can evaluate his M -best channels. In modern devices, multichannel sensing has become common practice that can be easily achieved by a beacon sent from the base station or other standard estimation techniques. Our approach does not require users to transmit this entire information to the base station.
- 3) Channel gains are quasi-static, such that the coherence time is larger than the time it requires to feed back the partial CSI (i.e., block-fading). This is also a very common assumption [10]–[14], [25], [26], since the multiuser diversity can only be exploited only in scenarios where this assumption holds.

We assume that the transmission power is the same in each channel, so no water-filling technique is used. This assumption is later justified both analytically and in simulations that show that the water-filling gain is negligible in our model for all practical SNR values. This has been verified for many similar scenarios [18]–[23].

Denote by \mathcal{A}_n the set of allocated channels of user n . The achievable rates of the users in bits per channel use, $\{R_n\}$, are defined as

$$R_n(\mathcal{A}_n) = \sum_{k \in \mathcal{A}_n} \log_2 \left(1 + \frac{g_{n,k}^2 P_{\max}}{bN_0} \right) \quad (1)$$

where P_{\max} is the maximal transmission power of each user and N_0 is the variance of the Gaussian noise in his receiver. Often, the goal of a channel allocation scheme is to maximize the sum-rate of the users

$$\begin{aligned} \max_{(\mathcal{A}_1, \dots, \mathcal{A}_N)} \sum_{n=1}^N R_n(\mathcal{A}_n) \\ s.t. \end{aligned} \quad (2)$$

$$\mathcal{A}_n \cap \mathcal{A}_m = \emptyset \quad \forall n, m, n \neq m$$

where the constraint means that each channel is allocated to a single user (orthogonal transmission). However, maximizing the sum-rate often comes at the expense of fairness.

$$\begin{aligned} \max_{(\mathcal{A}_1, \dots, \mathcal{A}_N)} \min_n R_n(\mathcal{A}_n) \\ s.t. \end{aligned} \quad (3)$$

$$\mathcal{A}_n \cap \mathcal{A}_m = \emptyset \quad \forall n, m, n \neq m$$

Our goal is to design a limited feedback channel allocation scheme that maximizes the sum-rate without compromising the minimal-rate (fairness).

A. Perfect Matching Allocation Scheme

In order for the BS to be able to compute the optimal solution of (2), all users need to feedback all their channel gains to the BS. Then, the optimal solution can be computed at the BS in polynomial time using the Hungarian Algorithm [32]. However, a feedback overhead of K quantized numbers for each user renders this approach impractical for cellular networks. Therefore, suboptimal schemes with limited feedback overhead and good performance are essential.

In general, there is a fundamental tradeoff between the amount of feedback used for the allocation and the performance achieved. By choosing a specific feedback structure, one chooses a specific operating point for this tradeoff. For example, by feeding back the values of the NK channel gains the optimal solution can be computed, and by feeding back one bit per channel per user (K bits per user), state of the art thresholding schemes can be applied [10]–[12].

We propose a novel limited feedback scheme that requires less feedback than current state of the art schemes. Not only that it does not perform worse, it even performs much better. The scheme is summarized in Algorithm 1. In our approach, each user transmits only the indices of his M -best channels, without their numerical value. Feeding back the indices of the M -best channels was proposed in [14] for an opportunistic channel allocation scheme. The novelty of our approach is not in the feedback structure but in the fact that our approach is non-opportunistic. This means that all channels and all users are simultaneously considered in the computation of the allocation, and not one by one. In particular, only in our scheme the results of the next section about the optimal M for a perfect matching are valid. Since [14] has nothing to do with perfect matchings, it suffers, like any opportunistic method, from an inherent non-fairness. We compare our performance to [14] in Section VI.

The information available at the BS using this limited feedback can be summarized using a bipartite graph defined as follows. Using this information, the BS aims to compute an allocation where each user gets b of his M -best channels. Such an allocation does not necessarily exist, and the existence probability depends on M .

Definition 1. A user-channel graph Γ is a balanced bipartite graph consisting of a user nodes set \mathfrak{N} and a channel nodes set \mathcal{K} . Every user $n \in \mathcal{N}$ has b user nodes (“agents”) and every channel has a single channel node. An edge is connected between $n_i \in \mathfrak{N}$ and $k \in \mathcal{K}$ if and only if channel k is one of the M -best channels for user n . A PM allocation corresponds to a PM in this graph.

From the perspective of the theory of random bipartite graphs, our graph is not typical. Each left side node has exactly M edges whereas each right side node may have any degree. Additionally, groups of b consecutive left side edges are connected to the exact same right side edges. In particular, this means that edge existence is not independent and that the graph is not

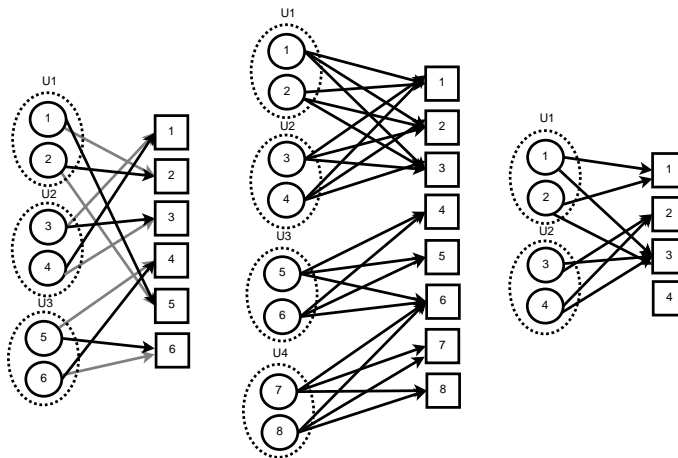


Fig. 2. User-channel graphs.

an Erdős–Rényi graph. This is why a novel proof for the existence of a PM is needed.

Examples of user-channel graphs can be seen in Figure 2. In the left graph $N = 3$, $b = 2$ and $M = 2$. A PM exists that consists of the grey edges. In the right most graph $N = 2$, $b = 2$ and $M = 2$. No user is connected to channel 4 so no PM can exist. In the central graph, $N = 8$, $b = 2$ and $M = 3$. Although each channel is connected to at least two user nodes, a PM does not exist. No matter which two channels are allocated to nodes 1 and 2, nodes 3 and 4 will be left with only one desired channel for both of them.

Algorithm 1 Perfect Matching Limited Feedback Scheme

Setting -There are N users in a cell and $K = bN$ channels for a positive integer b . Let $g_{n,k}$ be the channel gain of user n in channel k . Let $M = (1 + b)(1 + \varepsilon) \ln K$ for a positive ε .

1) **Periodically in the cell -**

- a) The base-station (BS) broadcasts a common beacon on all K channels.
 - b) Each user n independently -
 - i) Estimates the channel gains $g_{n,k}$ for each k .
 - ii) Transmits the subset of his M -best channels to the base-station, using $\log_2 \binom{K}{M}$ bits (see Algorithm 2).
 - c) Based on the information from (b).ii, the BS constructs the bipartite graph of Definition 1.
 - d) The BS computes the maximal matching in this graph (e.g., using the algorithm of [30]).
 - e) The BS transmits to each user the channels he is connected to in the matching from (d).
 - i) **Theorem 2 guarantees that with high probability this matching will assign each user b of his M -best channels (“perfect matching”).**
 - ii) **Theorem 8 guarantees that with high probability this allocation will have both asymptotically optimal sum-rate and max-min fairness.**
-

Algorithm 2 Feedback Encoding and Decoding

Encoding - The subset $\mathcal{S} \subset \{1, \dots, K\}$ is mapped to the index $E(\mathcal{S}, K)$, defined recursively as

- 1) If $\mathcal{S} = \emptyset$, then $E(\mathcal{S}, K) = 0$.
- 2) If $K \notin \mathcal{S}$, then $E(\mathcal{S}, K) = E(\mathcal{S}, K - 1)$.
- 3) If $K \in \mathcal{S}$, then $E(\mathcal{S}, K) = \binom{K-1}{|\mathcal{S}|} + E(\mathcal{S} \setminus \{K\}, K - 1)$.

Decoding - The index e is mapped to the subset $D(e, K, M)$, defined recursively as

- 1) If $e = 0$, then $D(0, K, M) = \emptyset$.
 - 2) If $e < \binom{K-1}{M}$ then $D(e, K, M) = D(e, K - 1, M)$.
 - 3) If $e \geq \binom{K-1}{M}$ then $D(e, K, M) = D\left(e - \binom{K-1}{M}, K - 1, M - 1\right) \cup \{K\}$.
-

TABLE I
NOTATIONS AND SYMBOLS USED THROUGH THIS PAPER.

N	Number of users
K	Number of channels
b	Number of channels per user
M	Number of best channels for each user.
$g_{n,k}$	Channel gain of user n in channel k
P_{\max}	maximal transmission power
$X_{(i)}$	The i -th smallest variable among X_1, \dots, X_N
\mathcal{E}_K	Probability that a PM does not exist with K channels

III. ASYMPTOTIC EXISTENCE OF PM ALLOCATIONS

Our goal is to show that a PM allocation, where each user only gets M -best channels, exists. There is a natural trade-off for the parameter M . A large value of M increases the probability that a PM allocation exists at the cost of reduced performance, since the M -best channel gets worse as M increases. The feedback communication overhead also increases with M . Hence, it is important to identify the smallest value of M for which a PM exists. In this section we characterize this optimal value of M that is used in Algorithm 1.

The main theorem of this paper is stated as follows:

Theorem 2 (Main Theorem). *Let $K = bN$ for a positive integer b that is constant with respect to N . Let \mathcal{E}_K be the event that a PM, where each user gets b channels that are among his M -best, does not exist. If $M \geq (b + 1)(1 + \varepsilon) \ln K$ for some $\varepsilon > 0$ then, for large enough K ,*

$$\Pr(\mathcal{E}_K) \leq \frac{1}{K^{\varepsilon + \frac{\varepsilon+1}{b}}} + \frac{\pi\sqrt{b}e^3}{12} \frac{1}{K^{\frac{3}{2}\varepsilon + \frac{3\varepsilon+1}{b}}} \quad (4)$$

Thus, a PM exists with a probability that approaches one as $K \rightarrow \infty$. Furthermore, if $\varepsilon \geq 1$, then the indicator of \mathcal{E}_K converges to zero almost surely.

Proof: See Appendix A. ■

For a PM to exist, every channel should be one of the M -best channels for at least one user. The following lemma shows that for this to occur, M must grow with respect to K (faster than $b \ln K$) and that $M = b(1 + \varepsilon) \ln K$ is enough for that purpose.

Lemma 3. *Let $K = bN$ for a positive integer b that is constant with respect to N . Let $M > 0$. Denote by \mathcal{E}_0 the event in which there is a channel which is not one of the M -best channels for any user.*

- 1) *If $M \leq b \ln K$ then $\lim_{K \rightarrow \infty} \Pr(\mathcal{E}_0) \geq \frac{1}{2}$.*
- 2) *If $M \geq b(1 + \varepsilon) \ln K$ and $\varepsilon > 0$ then $\lim_{K \rightarrow \infty} \Pr(\mathcal{E}_0) = 0$.*

Proof: See Appendix A. ■

IV. FEEDBACK REDUCTION

The implication of Theorem 2 is that our Algorithm 1 requires of each user to feedback only the indices of his $M = \lceil (b + 1)(1 + \varepsilon) \ln K \rceil$ best channels. This requires $\log_2 \binom{K}{M}$ bits from each user, to encode a subset of size M of a set of size K . Since $\log_2 \binom{K}{M} \leq M \log_2 K$, the amount of feedback bits scales like $O\left(\frac{\log_2^2 K}{K}\right)$, which improves with the number of channels K , therefore making our scheme especially appealing for large systems. In Algorithm 2 we present a practical method for encoding and decoding for indexing the subsets that requires exactly $\log_2 \binom{K}{M}$ bits, used for the combinatorial number system [33]. Its computational complexity is $O(K)$.

Proposition 4. *The encoding of Algorithm 2 requires $\log_2 \binom{K}{M}$ bits.*

Proof: The fact that $E(\mathcal{S}, K)$ is a bijection so each subset has a unique index and vice versa is proven in [34]. For each subset $\mathcal{S} \subset \{1, \dots, K\}$ we have

$$E(\mathcal{S}, K) \leq \sum_{m=0}^M \binom{K-1-m}{M-m} = \sum_{m=0}^M \binom{K-1-m}{K-1-M} \stackrel{(a)}{=} \binom{K}{M}$$

where (a) follows from the hockey-stick identity [34]. Hence, all the subsets $\mathcal{S} \subset \{1, \dots, K\}$ are mapped by $E(\mathcal{S}, K)$ to the indices 1 to $\binom{K}{M}$. ■

The feedback requirement of Algorithm 1 is compared to other schemes in Table 1 in bits per user per channel. Note that the optimal solution, given by the Hungarian algorithm, requires a quantization of q bits per user per channel, where a reasonable q that causes only a negligible performance degradation is $q = 8$.

State of the art limited feedback schemes [10]–[12] use a threshold for each channel gain. This means that each user has to send a packet of K data bits in each coherence time of the channel (several milliseconds). Already in LTE $K = 2048$ is a possibility [7], which results in a feedback (uplink) data rate of 409.6kbps for a coherence time of 5 milliseconds, more than four times the bandwidth of a voice call. In order to maintain multiuser diversity this feedback channel should be active very often. Besides creating a significant communication overhead in the uplink, this might not even be feasible due to battery constraints. Since the number of channels will increase in 5G systems, reducing the feedback amount is necessary if channel allocation with good performance is to be employed.

In general, in the high-SNR regime, the capacity of each user scales like $bB_W \ln(\ln N)$, where B_W is the bandwidth of a channel and $\ln(\ln N)$ is from the multiuser diversity in Rayleigh fading channels [13]. For a thresholding method the feedback rate scales like bN , so the feedback becomes even larger than the data rate in large systems. In [13], it was suggested to mitigate this overhead by grouping channels, at the cost of a reduced sum-rate (and minimal rate). Our method requires a feedback rate that scales like $\ln^2 N$ instead of like N as the thresholding schemes. Therefore, for modern systems with a relatively large B_W , the feedback overhead of our scheme remains negligible even for very large values of N . Hence, our method achieves the necessary feedback reduction, not only without reducing the performance, but actually improving

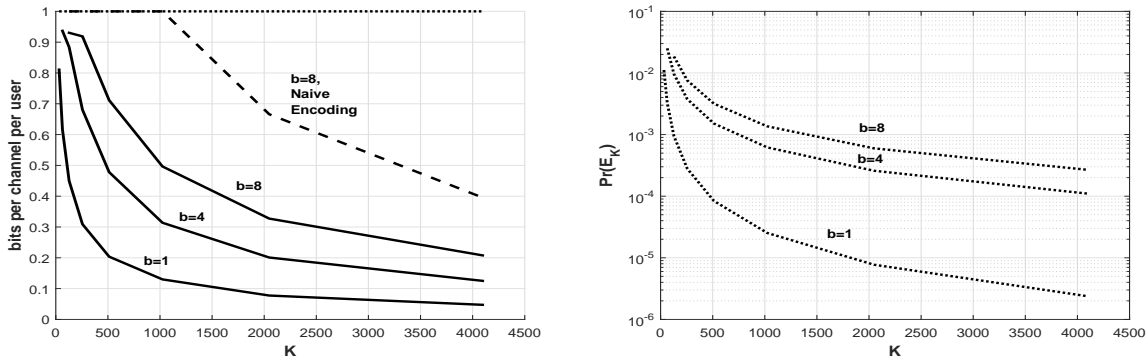


Fig. 3. Number of bits for feedback in Algorithm 1, and the corresponding non-perfect matching probability.

it. While the sum-rate improvement is moderate, the minimal rate is dramatically increased.

In Fig. 3 we present the number of feedback bits our algorithm requires as a function of the number of channels K , for $b = 1, 4, 8$ and $\varepsilon = 0.3, 0.7, 0.8$, respectively. We also show the number of bits required by a naive encoding scheme that directly transmits the M indices using $M \log_2 K = \log_2 K \lceil (b+1)(1+\varepsilon) \ln K \rceil$ bits. It is evident that our algorithm always requires less than the 1 bits per user per channel that state of the art thresholding approaches require [10]–[12]. For $N = 256$ and $K = 1024$, our approach requires less than 0.29 bits per user per channel - a saving of more than 70% in feedback overhead. For $K = 2048$ this becomes a saving of more than 80%. Even for small values $N = K = 32$, our approach requires less than 0.8 bits per user per channel. We also present the probability that a perfect matching does not exist, which is always less than 2% and improves very fast as K grows.

V. ASYMPTOTIC OPTIMALITY OF PM ALLOCATIONS

In the previous section we showed that $M = (b+1)(1+\varepsilon) \ln K$, for some $\varepsilon > 0$, is enough to guarantee that a PM allocation exists. In this section we show how good this PM allocation actually is. This has to do with the extent to which the M -best channel of each user is worse than his best channel. This definitely depends on the distribution of the channel gains. Luckily, all of the fading distributions used in practice tend to belong to the following class:

Definition 5 (Exponentially-Dominated Tail Distribution). Let X be a random variable with a CDF F_X . We say that X has an exponentially-dominated tail distribution if there exist $\alpha >$

$0, \beta \in \mathbb{R}, \lambda > 0, \gamma > 0$ such that

$$\lim_{x \rightarrow \infty} \frac{1 - F_X(x)}{\alpha x^\beta e^{-\lambda x^\gamma}} = 1 \quad (5)$$

The following simple proposition validates that the Rayleigh, m -Nakagami, Rice and Normal distributions have an exponentially-dominated tail.

Proposition 6. *The Rayleigh, m -Nakagami, Rice and Normal distributions are exponentially-dominated tail distributions.*

Proof: See Appendix A. ■

The fact that each user gets b of his M -best channels, and that M is kept relatively small for exponentially-dominated fadings, provides both sum-rate and fairness guarantees. This is a special property of our allocation. In general, maximizing the sum-rate does not provide any fairness guarantees for the users, as demonstrated in the following example.

Example 7. Consider the following $K \times K$ rates matrix for the case of $b = 1$, where $u_{n,k}$ is the rate of user n in channel k . Let $\delta > 0$.

$$\mathbf{U} = \begin{pmatrix} \delta & 1 & 1 & \cdots & 1 \\ 1 & 2 & 1 & \cdots & 1 \\ \delta & 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \delta & 1 & 1 & \cdots & 1 \end{pmatrix}$$

For this case, the Hungarian algorithm assigns channel k to user k for each k , resulting in a sum-rate of $K + \delta$ and a minimal rate of δ . Our PM approach assigns each user one of his M best-channels. If $2 \leq M \leq K - 1$, our algorithm results in a sum-rate of K and a minimal rate of 1. For a small enough δ , we obtain only a slight reduction in the sum-rate while having a dramatically better minimal rate.

The following theorem establishes the asymptotic optimality (in ratio) of a PM allocation both in the sum-rate and minimal rate senses, for exponentially-dominated tail fading distributions.

Theorem 8 (Sum Rate and Fairness Optimality Theorem). *Let the channel gains be the i.i.d.*

variables $g_{n,1}, \dots, g_{n,K}$, for each n . If, for each n and k , $g_{n,k}$ has an exponentially-dominated tail and $M = (b+1)(1+\varepsilon)\ln K$ for some $\varepsilon > 0$, then

$$\text{plim}_{K \rightarrow \infty} \frac{\min_{(\mathcal{A}_1, \dots, \mathcal{A}_N) \in \mathcal{P}} \sum_{n=1}^N R_n(\mathcal{A}_n)}{\max_{(\mathcal{A}_1, \dots, \mathcal{A}_N)} \sum_{n=1}^N R_n(\mathcal{A}_n)} = 1 \quad (6)$$

where \mathcal{P} is the set of PMs in the user-channel graph, and plim denotes convergence in probability.

Furthermore,

$$\text{plim}_{K \rightarrow \infty} \frac{\min_{(\mathcal{A}_1, \dots, \mathcal{A}_N) \in \mathcal{P}} \min_n R_n(\mathcal{A}_n)}{\max_{(\mathcal{A}_1, \dots, \mathcal{A}_N)} \max_n R_n(\mathcal{A}_n)} = 1. \quad (7)$$

Proof: Denote by $g_{n,(i)}$ the i -th smallest channel gain among $g_{n,1}, \dots, g_{n,K}$. We have

$$\begin{aligned} 1 &\geq \frac{\min_{(\mathcal{A}_1, \dots, \mathcal{A}_N) \in \mathcal{P}} \sum_{n=1}^N R_n(\mathcal{A}_n)}{\max_{(\mathcal{A}_1, \dots, \mathcal{A}_N)} \sum_{n=1}^N R_n(\mathcal{A}_n)} \geq \frac{b \sum_{n=1}^N \log_2 \left(1 + \frac{g_{n,(K-M+1)}^2 P_{\max}}{bN_0} \right)}{b \sum_{n=1}^N \log_2 \left(1 + \frac{g_{n,(K)}^2 P_{\max}}{bN_0} \right)} \geq \\ &\quad \frac{\min_n \log_2 \left(1 + \frac{g_{n,(K-M+1)}^2 P_{\max}}{bN_0} \right)}{\underbrace{\max_n \log_2 \left(1 + \frac{g_{n,(K)}^2 P_{\max}}{bN_0} \right)}_{A_K}} \geq \\ &\quad \frac{\log_2 \left(\left(1 + \frac{\max_n g_{n,(K)}^2 P_{\max}}{bN_0} \right) \frac{\min_n g_{n,(K-M+1)}^2}{\max_n g_{n,(K)}^2} \right)}{\log_2 \left(1 + \frac{\max_n g_{n,(K)}^2 P_{\max}}{bN_0} \right)} = 1 + \underbrace{\frac{2 \log_2 \left(\frac{\min_n g_{n,(K-M+1)}}{\max_n g_{n,(K)}} \right)}{\log_2 \left(1 + \frac{\max_n g_{n,(K)}^2 P_{\max}}{bN_0} \right)}}_{B_K}. \quad (8) \end{aligned}$$

The asymptotic behavior of $\frac{\min_n g_{n,(K-M+1)}}{\max_n g_{n,(K)}}$ is dictated by how fast the tail distribution function of $g_{n,k}$, defined as $\bar{F}_g(x) = 1 - F_g(x)$, vanishes to zero. For a $g_{n,k}$ with an exponentially-dominated tail distribution we have $\text{plim}_{K \rightarrow \infty} \log_2 \left(1 + \frac{g_{n,(K)}^2 P_{\max}}{N_0} \right) = \infty$ since the distribution has an infinite support, and also $\lim_{K \rightarrow \infty} \Pr \left(\frac{\min_n g_{n,(K-M+1)}}{\max_n g_{n,(K)}} > r \right) = 1$ for some $r > 0$. Therefore $\text{plim}_{K \rightarrow \infty} B_K = 0$ and (6),(7) are obtained from the Sandwich Theorem. Note that for $g_{n,k}$ with a bounded distribution we have $\text{plim}_{K \rightarrow \infty} \frac{\min_n g_{n,(K-M+1)}}{\max_n g_{n,(K)}} = 1$ since both the numerator and the denominator approach $F_g^{-1}(1) < \infty$, so $\text{plim}_{K \rightarrow \infty} B_K = 0$ still holds.

In [35], we find an upper bound U_K such that $\lim_{K \rightarrow \infty} \Pr \left(\max_n g_{n,(K)} \leq U_K \right) = 1$ and a lower bound L_K such that $\lim_{K \rightarrow \infty} \Pr \left(\min_n g_{n,(K-M+1)} \geq L_K \right) = 1$, and show that $\lim_{K \rightarrow \infty} \frac{L_K}{U_K} > 0$. This

establishes that indeed $\lim_{K \rightarrow \infty} \Pr \left(\frac{\min_n g_{n,(K-M+1)}}{\max_n g_{n,(K)}} > r \right) = 1$ for some $r > 0$. Specifically, Theorem V.11 in [35], shows that $\text{plim}_{K \rightarrow \infty} A_K = 1$, provided that $g_{n,1}, \dots, g_{n,K}$ have an exponentially-dominated tail. Its proof follows the above arguments. ■

Theorem 8 has a stronger argument than simply asymptotic optimality of the sum-rate and the minimal rate. Observe that the denominator in (7) consists of the best rate a user can receive in any allocation, even an allocation that only favors him. This rate is achieved when user n is allocated his best b channels from out of the entire K channels. Hence, even the user with the minimal rate, in all possible PM allocations, achieves asymptotically full multiuser diversity. In fact, this is also true for the bN agents as well, which leads to the following corollary.

Corollary 9. *Let $K = bN$ for a positive integer b that is constant with respect to N . The ratio between the rate achieved by water-filling to that of an equal power allocation converges in probability to one as $K \rightarrow \infty$.*

Proof: Denote by $\{P_{n,k}^*\}$ the set of the optimal transmission powers given $\{g_{n,k}\}$. We have

$$\frac{\sum_{k \in \mathcal{A}_n} \log_2 \left(1 + \frac{g_{n,k}^2 P_{\max}}{bN_0} \right)}{\sum_{k \in \mathcal{A}_n} \log_2 \left(1 + \frac{g_{n,k}^2 P_{n,k}^*}{N_0} \right)} \stackrel{(a)}{\geq} \frac{\sum_{k \in \mathcal{A}_n} \log_2 \left(1 + \frac{g_{n,(K-M+1)}^2 P_{\max}}{bN_0} \right)}{\sum_{k \in \mathcal{A}_n} \log_2 \left(1 + \frac{g_{n,(K)}^2 P_{\max}}{bN_0} \right)} \geq \frac{\min_n \log_2 \left(1 + \frac{g_{n,(K-M+1)}^2 P_{\max}}{bN_0} \right)}{\max_n \log_2 \left(1 + \frac{g_{n,(K)}^2 P_{\max}}{bN_0} \right)} \quad (9)$$

If $g_{n,k}$ is increased to $g_{n,(K)}$ for each k and the transmission powers are kept as $P_{n,k}^*$ for each k , then the sum-rate increases. If for the increased channel gains the optimal power allocation is used instead of $\{P_{n,k}^*\}$, the sum-rate will further increase. The optimal power allocation for K identical channel gains of $g_{n,(K)}$ is an equal power allocation. These arguments yield (a), together with $g_{n,k}^2 \geq g_{n,(K-M+1)}^2$ for each k , which holds in a PM allocation with parameter M . ■

VI. SIMULATION RESULTS

In this section, we demonstrate our analytical results using numerical simulations. We also validate the fact that our findings hold in a broader model that includes correlated channels and an allocation of unequal numbers of channels to users. All the rates are measured in bits per second, assuming that the bandwidth of a channel is 15kHz (as in LTE [36]). We used

a Rayleigh fading network; i.e., $\{g_{n,k}\}$ are Rayleigh distributed and $\{g_{n,k}^2\}$ are exponentially distributed. In each experiment, we ran 100 random realizations of the channel gains. Unless otherwise states, the transmission powers were chosen such that the mean SNR for each link was 20dB. We used the Hungarian algorithm [32] to compute the optimal sum-rate solution. In order to compute the maximal matchings in the bipartite graph (which Theorem 3 guarantees are perfect matchings), we used the algorithm in [28]. This algorithm requires a time complexity of $O(K \log_2^2 K)$ instead of $O(K^3)$ for the Hungarian algorithm, which also could have been used for the purpose of computing a maximal matching (on the binary preferences matrix).

A. Uncorrelated Identically-Distributed Channels

In this subsection, the channel gains were uncorrelated with parameter $\lambda = 1$ for all exponential variables. We used the parameter $M = \lceil 1.5(b+1) \ln(K) \rceil$ and ran the simulations for $N = 10, 25, 50, 75, 100$. We found that for every realization a PM existed. This implies that our asymptotic results are valid for as low as $N = 10$.

In Figure 4, the mean and minimal rates are presented as a function of N for uncorrelated channels. We used $b = 4$. It is evident that the performance of the PM allocation was already above 90% of the optimal allocation for $N = 50$. As anticipated by our results, it improved with N . We can also see that the minimal rate of the optimal allocation was good, which was to be expected. The chance of a single user getting $b = 4$ bad channels is smaller than in the case of $b = 1$. The random allocation rates, which can be thought of as the result of an allocation that ignores the CSI, were far behind. This effect is crucial especially for the minimal rate. In [12], the base station assigns each channel to one of the users that exceeded a certain threshold (possibly with defined priorities between users). In order to make the algorithm competitive in the minimal rate sense as well, each time multiple users exceeded the threshold, we allocated the channel to the user with the minimal number of channels thus far. We used $b = 4$, so $K = 4N$. Our mean rate is slightly better than that of [12] for all $N \geq 20$. The minimal rate of [12] is much lower ($\sim 78\%$ of our minimal rate), and does not increase with N while our minimal rate does. Last but not least, we require $\frac{1}{K} \log_2 \binom{K}{M}$ bits per user channel (0.81 for $N = 10$ and 0.52 for $N = 200$), while [12] always requires one bit per user per channel (exceeded

the threshold or not). Hence, our algorithm outperforms [12] both in terms of performance and feedback communication overhead.

In Figure 5, we compare our algorithm also to that of [14], which has two versions. This algorithm assigns a single channel to each user, so we had choose $b = 1$. We used the same $M = \lceil 3 \ln(K) \rceil$ for our algorithm and that of [14]. The mean rate of the set-best version is very similar to ours, with exactly the same feedback communication overhead. The mean-rate in the ordered-best version is a little better than our mean-rate, but has a significant feedback cost. For $N = K = 20$, the ordered-best requires ~ 2.25 bits per user per channel, while the set-best (and our approach) requires ~ 0.87 bits per user per channel. For $N = K = 100$, the ordered-best requires ~ 0.98 bits per user per channel, while the set-best and our approach require ~ 0.55 bit per user per channel. However, our major anticipated advantage is in the minimal rate. The minimal rate of [14] is very close in both versions and is only about $\sim 70\%$ of our minimal rate, and decreases with N while our minimal rate increases. The minimal rate of [12] converges to zero, since for $N = K$, the probability that some users will not be allocated a channel is high.

To show the tightness of our selected M we now demonstrate the threshold phenomena for the existence of a PM. We reduced M to $M = \lceil b \ln(K) \rceil$ instead of $M = \lceil 1.5(b+1) \ln(K) \rceil$, with $b = 4$. As presented in Table II, the probability of a PM was at least 0.45 already from $N = 10$. These results are in agreement with Lemma 3, that shows that for this value of M and the probability for a PM is bounded from below by “ $0.5 - \delta$ ” for large enough N (and arbitrarily small δ).

In Figure 6, we present the rates as a function of b for $N = 30$. All the rates increased almost linearly with b , and the ratio of the rates of a PM allocation to the optimal allocation rates remained almost constant for all b . This implies that the results of Figure 4 are similar for each b value.

TABLE II
EXISTENCE OF PM ALLOCATIONS AS A FUNCTION OF N FOR $M = \lceil b \ln(K) \rceil$ AND $b = 4$.

N	10	25	50	75	100
$\Pr(\mathcal{E}_{bN})$	0.56	0.45	0.64	0.59	0.52

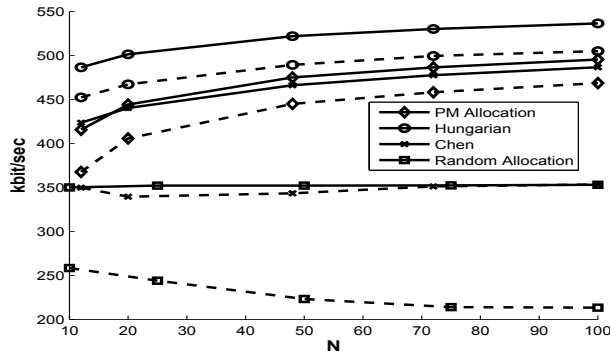


Fig. 4. Achievable rates as a function of N for uncorrelated channels. Solid lines represent the mean rates and dashed lines the minimal rates.

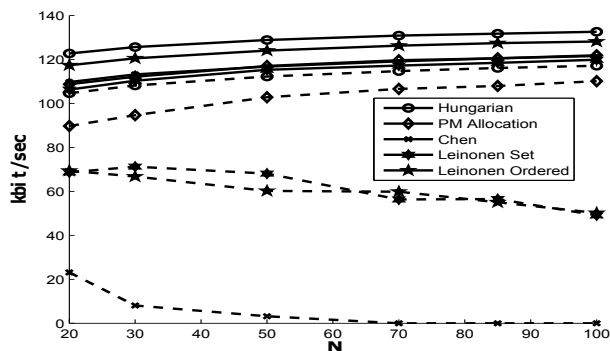


Fig. 5. Achievable rates as a function of N compared to Leinonen's Algorithm, Chen's Algorithm and the Hungarian Algorithm. Solid lines represent the mean rates and dashed lines the minimal rates.

B. Correlated Non-Identically Distributed Channels

In this subsection, we used correlated resource blocks as the channels, each consisting of 12 subcarriers. The rates are normalized per subcarrier. We used the Extended Pedestrian A model (EPA, see [36]) for the excess tap delay and the relative power of each tap. In this model, the channel gains of different users are not identically distributed, which is the practical scenario. We used $N = 32, 64, 96, 128$.

Figure 7 shows the mean and minimal rates as a function of N averaged over 100 realizations and with $b = 4$. The essential difference with correlated channels is that the diversity decreases; namely, adjacent channels are likely to be good or bad together. This makes the multiuser diversity gain lower and as a result the rates are lower, and look like the rates of the uncorrelated fading for smaller values of N . Additionally, this reduced diversity lead us to choose $M = \lceil 2(b+1)\ln(K) \rceil$, since a higher M is needed for a PM to exist in the correlated user-channel

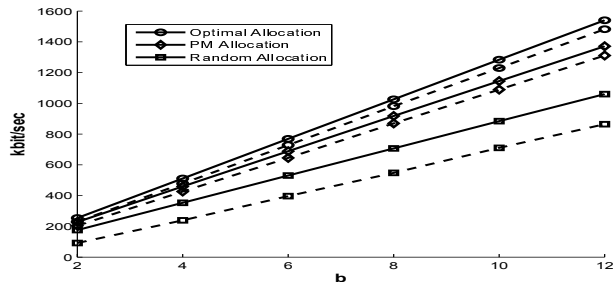


Fig. 6. Achievable rates as a function of b for $N = 20$. Solid lines are used for the mean rates and dashed lines for the minimal rates.

graph. Indeed, we always found PMs for this choice of M . However, a larger M makes the PM allocation rates lower. Nevertheless, these rates are still more than 85% of optimal rates already for $N = 50$ and get closer as N increases. Given these experimental results, we conjecture that our analytical results can be extended to the case of correlated channels

In Figure 8, we present the rates for correlated resource blocks the with $b = 1$. Note that both [14] and [12] assumed uncorrelated channels and identically distributed channel gains. In a thresholding approach, a different threshold can be used for each user, compromising mean-rate for fairness. However, since the minimal rate of [12] converges to zero anyway, we chose the threshold in order to optimize the mean-rate. As shown, the results are similar those of Figure 4, and the feedback overhead comparison is of course the same. The main difference is that the minimal rate of all methods is lower. This does not represent a degradation for any of the algorithms, since with non-identically distributed channel gains, the minimal rate user is the user with the worst global conditions. In our method, the minimal rate is still significantly better than those of [14] and [12], since this minimal rate user got one of his best channels (even though all of them were relatively bad).

C. Unequal Channel Allocation

In this subsection, we allocated users a different number of channels, by dividing them into four classes. Specifically, $\frac{N}{4}$ of the users were allocated a single channel, $\frac{N}{4}$ two channels, $\frac{N}{4}$ three channels and $\frac{N}{4}$ four channels. We used $M_i = \lceil 3.75 \ln(K) \rceil$ for all users, which amounts to using the average b in $M = \lceil 1.5(b + 1) \ln(K) \rceil$. Such a differentiation might occur if different users require a different quality of service (QoS). We simulated 100 realizations for each value

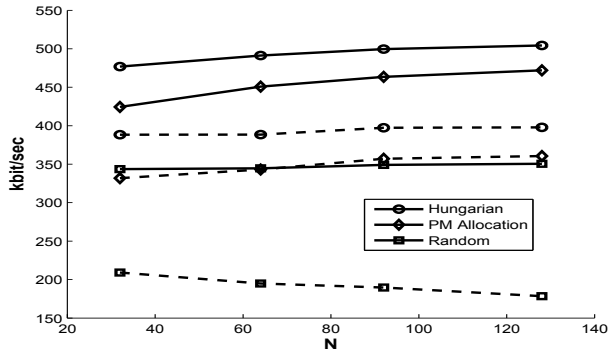


Fig. 7. Achievable rates as a function of N for correlated channels. Solid lines represent the mean rates and dashed lines the minimal rates.

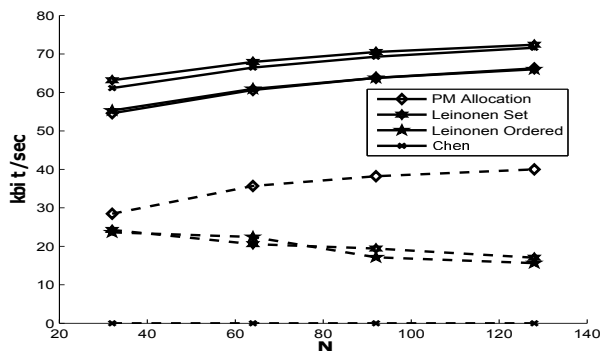


Fig. 8. Achievable rates as a function of N for non-identically distributed correlated channels, compared to state of the art algorithms. Solid lines represent the mean rates and dashed lines the minimal rates.

from $N = 12, 20, 48, 72, 100$. PMs still existed in all realizations. We ran another 100 realizations with correlated resource blocks, in the same manner as in Figure 7 and used $M = \lceil 5 \ln(K) \rceil$ for all users. Figure 9 shows the empirical CDFs of the rates for the four classes, for uncorrelated channels and correlated resource blocks, with $N = 100$ and $N = 128$. Note that the PM allocation rates are quite concentrated and most of the users get similar rates. These rates are always better than those of a random allocation, and close the optimal allocation rates. These results confirm our belief that our analysis can be generalized to the case of an unequal allocation, which is further supported by our proof itself.

D. Water-Filling Gain

In all the above results, water-filling over the b channels of each user had a negligible gain in terms of kbps. To be exact, it never improved any of the rates by more than 0.01 kbps. This is consistent with our analytical result in Corollary 9. We repeated our experiment for an SNR

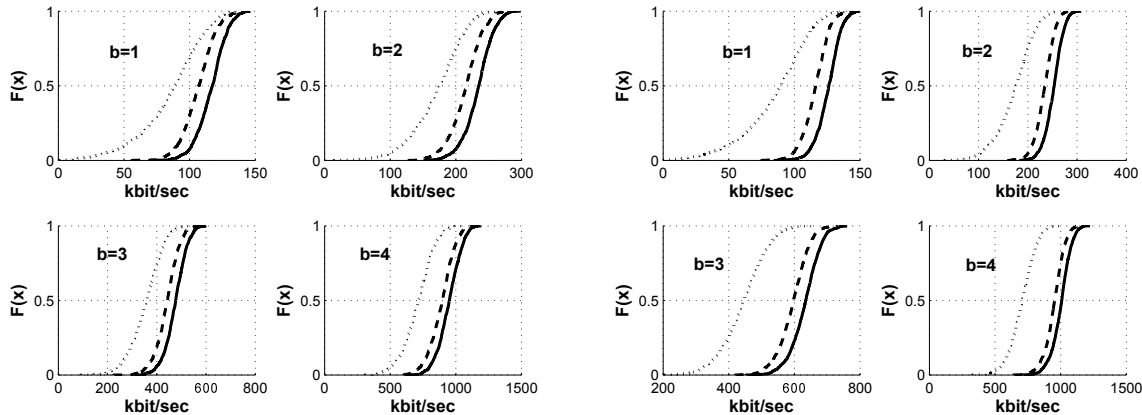


Fig. 9. Empirical CDFs of the rates for different classes of users with $N = 100$, for uncorrelated channels (left) and correlated resource blocks (right). Solid lines are used for optimal rates, dashed lines for the PM allocation rates and dotted lines for the random allocation rates.

of 10dB, $N = 30$ and $M = \lceil 2(b+1) \ln(K) \rceil$, where the water-filling gain should be more significant than with our original parameters. We used $b = 2, 4, 6, 8, 10, 12, 14$. In Figure 10 we present the water filling relative mean-rate gain, defined as $\frac{R_{WF}-R}{R}$, as a function of b . Clearly, even with SNR of 10dB, water-filling did not improve the mean rates by more than 0.16% for either the Hungarian algorithm or our PM allocation for any value of b .

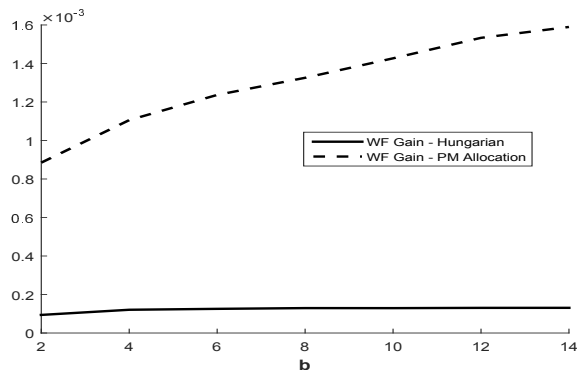


Fig. 10. The water-filling relative mean-rate gain $\frac{R_{WF}-R}{R}$ as a function of b for $N = 30$.

VII. CONCLUSIONS

In this paper we suggested an approach aimed at achieving the multiuser diversity gain for resource block allocation under the frequency-selective channel. Our approach requires knowledge of the indices of the M -best channels for each of the N users, instead of the whole NK channel gains.

We showed that $M = (b + 1)(1 + \varepsilon) \ln(K)$ for a positive ε is sufficient to guarantee that a PM allocation, where each user gets all his b channels from his M -best channels, asymptotically exists. For this value of M , our suboptimal approach was shown to be asymptotically optimal for a broad class of fading distributions, both in the sum-rate and max-min senses.

Our algorithm is the first limited feedback algorithm that has provably asymptotic optimality both of the sum-rate and the minimal rate. It also requires significantly less feedback than the state of the art algorithms that achieve a good sum-rate. Hence, it constitutes a major step toward achieving the multiuser diversity gain of a selective channel in practice.

A future direction could involve trying to lift the assumption of i.i.d. channel gains by only assuming m -dependent channel gains and using the techniques presented in [37].

APPENDIX A

PROOFS

For our results, we need the following two lemmas.

Lemma 10. For each $0 \leq l \leq K$, we have $\frac{\binom{l}{M}}{\binom{K}{M}} \leq \left(\frac{l}{K}\right)^M$.

Proof:

$$\frac{\binom{l}{M}}{\binom{K}{M}} = \frac{\frac{l!}{(l-M)!}}{\frac{K!}{(K-M)!}} = \frac{\prod_{i=1}^M (l - M + i)}{\prod_{i=1}^M (K - M + i)} = \prod_{i=1}^M \frac{l - M + i}{K - M + i} = \prod_{j=0}^{M-1} \frac{l - j}{K - j} \stackrel{(a)}{\leq} \left(\frac{l}{K}\right)^M \quad (10)$$

where (a) comes from the fact that if $l < K$ then $f(x) = \frac{l-x}{K-x}$ is monotonically decreasing for all $x < K$, so $\frac{l-j}{K-j} \leq \frac{l}{K}$ for each $0 \leq j < M$. ■

Lemma 11. Let $M \geq (b + 1)(1 + \varepsilon) \ln K$ for some $\varepsilon > 0$. Define the function

$$f(x) = \ln(2) K \left(1 + \frac{1}{b}\right) h_b\left(\frac{x+1}{K}\right) - \frac{M}{b} \left(1 - \frac{1}{K}\right) x + \frac{1}{b} \frac{M}{K} x^2 - \frac{M}{b} \quad (11)$$

where h_b is the binary entropy function (see [38, page 666]). Then, for large enough K , we have for all $M \leq x \leq K - 1$

$$f(x) \leq 3 - \left(1 + \frac{1}{b}\right) \left(1 + \frac{3}{2}\varepsilon\right) \ln(K) \quad (12)$$

Proof: Since $\frac{dKh_b\left(\frac{x+1}{K}\right)}{dx} = \log_2\left(\frac{K}{x+1} - 1\right)$ we have

$$\frac{df(x)}{dx} = \left(1 + \frac{1}{b}\right) \ln\left(\frac{K}{x+1} - 1\right) - \frac{M}{b} \left(1 - \frac{1}{K}\right) + \frac{2}{b}x \frac{M}{K} \quad (13)$$

and

$$\frac{d^2f(x)}{dx^2} = -\frac{1 + \frac{1}{b}}{K - x - 1} - \frac{1 + \frac{1}{b}}{x + 1} + \frac{2M}{bK}. \quad (14)$$

so for any $x \in \left[\frac{(b+1)K}{M} - 1, K - \frac{(b+1)K}{M} - 1\right]$ we have $\frac{d^2f(x)}{dx^2} \geq 0$. If $x \geq K - \frac{(b+1)K}{M} - 1$ then for $M \geq (b+1)(1+\varepsilon)\ln(K)$ we get, for large enough K that

$$\frac{df(x)}{dx} \geq \left(1 - \frac{2b+2}{M} - \frac{1}{K}\right) \frac{M}{b} - \left(1 + \frac{1}{b}\right) \ln(K-1) \geq 0 \quad (15)$$

also if $M < x \leq \frac{(b+1)K}{M} - 1$ then for $M \geq (b+1)(1+\varepsilon)\ln(K)$ we get from (13), for large enough K

$$\frac{df(x)}{dx} \leq \left(1 + \frac{1}{b}\right) \ln\left(\frac{K}{M} - 1\right) - \frac{M}{b} \left(1 - \frac{1}{K}\right) + 1 + \frac{2}{b} \leq 0. \quad (16)$$

So $f(x)$ is convex in $\left[\frac{(b+1)K}{M} - 1, K - \frac{(b+1)K}{M} - 1\right]$, increasing in $\left[K - \frac{(b+1)K}{M} - 1, K - 2\right]$ and decreasing in $\left[M, \frac{(b+1)K}{M} - 1\right]$. We conclude that the maximum of $f(x)$ must be attained at the borders $l = M, K - 2$.

Observe that for all $l < \frac{K}{2}$

$$Kh_b\left(\frac{l}{K}\right) = l \log_2 \frac{K}{l} - (K-l) \log_2 \left(1 - \frac{l}{K}\right) \stackrel{(a)}{\leq} 2l + l \log_2 \frac{K}{l} \quad (17)$$

where in (a) we used $\log_2\left(1 - \frac{l}{K}\right) \geq -\frac{2l}{K}$ which is true for $l \leq \frac{K}{2}$. Now we substitute $x =$

$M, K - 2$ in $f(x)$ to obtain

$$\begin{aligned} f(M) &= \ln(2) \left(1 + \frac{1}{b}\right) Kh_b \left(\frac{M+1}{K}\right) - \frac{M^2}{b} \left(1 - \frac{1}{K}\right) + \frac{M^3}{bK} - \frac{M}{b} \\ &\stackrel{(a)}{\leq} \left(1 + \frac{1}{b}\right) \left(2 \ln(2) M + (M+1) \ln \frac{K}{M+1}\right) - \frac{M^2}{b} \left(1 - \frac{1}{K}\right) + \frac{M^3}{bK} \stackrel{(b)}{\leq} \\ &\quad M \left(1 + \frac{1}{b}\right) \left(3 + \left(1 + \frac{1}{M}\right) \ln K - \frac{M}{b+1} \left(1 - \frac{1}{K}\right)\right) \end{aligned} \quad (18)$$

where (a) is from (17) and (b) is for large enough K due to $M \geq (b+1)(1+\varepsilon)\ln(K)$. Next

$$\begin{aligned} f(K-2) &= \ln(2) \left(1 + \frac{1}{b}\right) Kh_b \left(\frac{1}{K}\right) - \frac{M(K-1)(K-2)}{bK} + \frac{M(K-2)^2}{bK} - \frac{M}{b} \stackrel{(a)}{\leq} \\ &\quad \left(1 + \frac{1}{b}\right) (\ln K + 2 \ln(2)) - \frac{2M}{b} \left(1 - \frac{1}{K}\right) \end{aligned} \quad (19)$$

where (a) is from $h_b(1 - \frac{1}{K}) = h_b(\frac{1}{K})$. Hence for large enough K (19) is larger than that in (18). We conclude that for each $l = 1, \dots, K-2$

$$\begin{aligned} f(x) \leq \max_x f(x) &\leq \left(1 + \frac{1}{b}\right) (\ln K + 2 \ln(2)) - \frac{2M}{b} \left(1 - \frac{1}{K}\right) \leq \\ &3 - \left(1 + \frac{1}{b}\right) \ln(K) \left(2(1+\varepsilon) \left(1 - \frac{1}{K}\right) - 1\right) \leq \\ &\stackrel{(a)}{\leq} 3 - \left(1 + \frac{1}{b}\right) \left(1 + \frac{3}{2}\varepsilon\right) \ln(K) \end{aligned} \quad (20)$$

where (a) is true for large enough K such that $\frac{1}{K} \leq \frac{\varepsilon}{2}$. ■

A. Proof of Theorem 2

Proof: The bipartite graph can also be represented using a $K \times K$ preferences matrix U . Our matrix U is a simplified binary matrix where $u_{n,k} = 0$ if channel k is one of the M -best channels of agent n and $u_{n,k} = 1$ otherwise. It consists of $\frac{K}{b}$ submatrices of b identical rows each, and exactly M zeros at each row. We want to show that the probability that the Hungarian algorithm on U converges after one iteration approaches one as $K \rightarrow \infty$; hence the resulting cost is zero, which represents a PM in the equivalent bipartite graph. Note that since $K = bN$ for a positive integer b , when $K \rightarrow \infty$ then also $N \rightarrow \infty$. Denote by $Q_l(K, M)$ the probability that U can be covered by $K - l - 1$ rows and l columns such that all its zeros are covered, and that there is no smaller l with this property. We want to upper bound $Q_l(K, M)$ for each

l . Now observe that if U is covered such that a row is covered but one of its identical rows is not, this means that U could be covered without this row cover. Hence, if the row covers do not include all the identical rows, they are redundant. This means that only values of $K - l - 1$ (or $l + 1$) that are a multiple of b should be considered, and $Q_l(K, M) = 0$ for all other values. We proceed by counting the number of combinations of each such matrix

$$Q_l(K, M) \leq \underbrace{\binom{K}{l}}_A \underbrace{\binom{\frac{K}{b}}{\frac{l+1}{b}} \binom{K}{M}}_B \underbrace{\binom{l}{M}}_C \binom{K}{M}^{-\frac{K}{b}} = \binom{K}{l} \binom{\frac{K}{b}}{\frac{l+1}{b}} \binom{l}{M}^{\frac{l+1}{b}} \binom{K}{M}^{-\frac{l+1}{b}} \quad (21)$$

where A is the number of combinations for the index of the $K - l - 1$ covered rows and the l covered columns. Now we have to count the combinations for the structure of the covered and uncovered rows. The factor B is the number of combinations for $K - l - 1$ covered rows with exactly M zeros in each. We are left with $l + 1$ rows that we know can be covered by l columns, and no less than l . Hence each of these rows has zeros only in these l columns, so there can be no more than C such rows. Finally, the denominator $\binom{K}{M}^{\frac{K}{b}}$ is the number of combinations of a matrix with M zeros in each row, that constitute the whole probability space.

For $l = K - 1$ we get

$$Q_{K-1}(K, M) \stackrel{(a)}{\leq} K \left(1 - \frac{1}{K}\right)^{\frac{M}{b}K} \stackrel{(b)}{\leq} K e^{-\frac{M}{b}} \quad (22)$$

where (a) follows from Lemma 10 and (b) from the inequality $(1 - \frac{1}{K})^K \leq e^{-1}$. Note that for $l = 1, \dots, M - 1$ it is trivial that $Q_l(K, M) = 0$ since each row has exactly M zeros. For the rest of the proof we assume that $M < l \leq K - 2$. For these values, we have

$$\binom{K}{l} \binom{\frac{K}{b}}{\frac{l+1}{b}} = \frac{l+1}{K-l} \binom{K}{l+1} \binom{\frac{K}{b}}{\frac{l+1}{b}} \stackrel{(a)}{\leq} \frac{\sqrt{b}}{2\pi} \frac{1}{(K-l) \left(1 - \frac{l+1}{K}\right)} 2^{K(1+\frac{1}{b})h_b(\frac{l+1}{K})} \quad (23)$$

where in (a) we used $\binom{K}{l} \leq \sqrt{\frac{1}{2\pi l(1-\frac{l}{K})}} 2^{Kh_b(\frac{l}{K})}$, which is valid for all $0 < l < K$ (where

h_b is the binary entropy function, see [38, page 666]). Combining the above we obtain from (21) the following upper bound

$$Q_l(K, M) \stackrel{(a)}{\leq} \frac{\sqrt{b}}{2\pi(K-l)\left(1-\frac{l+1}{K}\right)} 2^{K\left(1+\frac{1}{b}\right)h_b\left(\frac{l+1}{K}\right)} \left(\frac{l}{K}\right)^{\frac{M}{b}(l+1)} \stackrel{(b)}{\leq} \frac{\sqrt{b}}{2\pi(K-l)\left(1-\frac{l+1}{K}\right)} e^{\ln(2)K\left(1+\frac{1}{b}\right)h_b\left(\frac{l+1}{K}\right) - \frac{M}{Kb}(K-l)(l+1)} \quad (24)$$

where (a) is due to Lemma 10 and (23) and (b) due to $1 - \frac{K-l}{K} \leq e^{-\frac{K-l}{K}}$. We want to find the maximal value of this bound as a function of l . Define the auxiliary function

$$f(x) = \ln(2)K\left(1+\frac{1}{b}\right)h_b\left(\frac{x+1}{K}\right) - \frac{M}{b}\left(1-\frac{1}{K}\right)x + \frac{1}{b}\frac{M}{K}x^2 - \frac{M}{b} \quad (25)$$

and denote by \mathcal{E}_K the event in which a PM does not exist. We conclude by the union bound that

$$\begin{aligned} \Pr(\mathcal{E}_K) &\leq \\ Q_{K-1}(KM) + \sum_{l=M}^{K-2} Q_l(K, M) &\leq \frac{1}{K^\varepsilon} + Ke^{-\frac{M}{b}} + \sum_{l=M}^{K-2} \frac{\sqrt{b}K}{2\pi(K-l)(K-l-1)} e^{f(l)} \stackrel{(a)}{\leq} \\ Ke^{-\frac{M}{b}} + \frac{\sqrt{b}K}{2\pi} \frac{e^3}{K^{\frac{3}{2}\varepsilon + \frac{3\varepsilon+1}{b}}} \sum_{l=M}^{K-2} \frac{1}{(K-l-1)^2} &\stackrel{(a)}{\leq} Ke^{-\frac{M}{b}} + \frac{\sqrt{b}K}{2\pi} \frac{e^3}{K^{\frac{3}{2}\varepsilon + \frac{3\varepsilon+1}{b}}} \sum_{k=1}^{\infty} \frac{1}{k^2} \leq \\ &\frac{1}{K^{\varepsilon + \frac{\varepsilon+1}{b}}} + \frac{\pi\sqrt{b}e^3}{12} \frac{1}{K^{\frac{3}{2}\varepsilon + \frac{3\varepsilon+1}{b}}} \quad (26) \end{aligned}$$

where (a) follows from Lemma 11 for large enough K , and (b) follows by substituting $k = K-l$ and adding terms to the sum. Note that if $\varepsilon \geq 1$ then $\sum_{K=1}^{\infty} \Pr(\mathcal{E}_K) < \infty$. Hence, by the Borel-Cantelli lemma, the probability that an infinite number of events from $\{\mathcal{E}_K\}$ will occur is zero. Hence, we conclude that the indicator of \mathcal{E}_K converges almost surely to zero. \blacksquare

B. Proof of Lemma 3

Proof: Denote by $\mathcal{E}_{0,k}$ the event in which channel k is not one of the M -best channels for any of the users. We want to bound from below and above the probability that there exists a channel that is not good for any of the users; i.e., the probability of $\mathcal{E}_0 = \bigcup_k \mathcal{E}_{0,k}$. Due to the i.i.d. assumption on the channel gains and their independency between users we have

$\Pr(\mathcal{E}_{0,k}) = \left(1 - \frac{M}{K}\right)^N$. From the union bound we obtain that

$$\Pr(\mathcal{E}_0) = \Pr\left(\bigcup_k \mathcal{E}_{0,k}\right) \leq K \left(1 - \frac{M}{K}\right)^N. \quad (27)$$

Since $\left(1 - \frac{M}{K}\right)^N$ decreases with M , using $M \geq b(1 + \varepsilon) \ln K$ leads to

$$\lim_{K \rightarrow \infty} \Pr(\mathcal{E}_0) = \lim_{K \rightarrow \infty} K \left(1 - \frac{M}{K}\right)^N \leq \lim_{K \rightarrow \infty} K \left(1 - \frac{b(1 + \varepsilon) \ln K}{K}\right)^{\frac{K}{b}} = \lim_{K \rightarrow \infty} K e^{-(1 + \varepsilon) \ln K} = 0 \quad (28)$$

Now we want to bound from below the probability of a PM. We do so by using the inclusion-exclusion principle. We use the following lower bound

$$\Pr\left(\bigcup_k \mathcal{E}_{0,k}\right) \geq \sum_k \Pr(\mathcal{E}_{0,k}) - \sum_{k_1, k_2} \Pr(\mathcal{E}_{0,k_1} \cap \mathcal{E}_{0,k_2}). \quad (29)$$

By direct counting of the number of user-channel graphs with two undesired channels we obtain

$$\sum_{k_1, k_2} \Pr(\mathcal{E}_{0,k_1} \cap \mathcal{E}_{0,k_2}) = \binom{K}{2} \binom{K-2}{M}^N \binom{K}{M}^{-N} \stackrel{(a)}{\leq} \frac{K^2}{2} \left(1 - \frac{2}{K}\right)^{M \frac{K}{b}} \quad (30)$$

Since $\binom{K}{2}$ is the number of choices for these two channels, $\binom{K-2}{M}$ is the number of combinations of the edges of each agent and $\binom{K}{M}^N$ is the total number of user-channel graphs. Inequality (a) is from Lemma 10. We obtain

$$\Pr(\mathcal{E}_0) = \Pr\left(\bigcup_k \mathcal{E}_{0,k}\right) \geq K \left(1 - \frac{M}{K}\right)^N - \frac{K^2}{2} \left(1 - \frac{2}{K}\right)^{M \frac{K}{b}} \quad (31)$$

Obviously, $\Pr(\mathcal{E}_{0,k})$ decreases with M . Hence, using $M = b \ln K$, we conclude that

$$\lim_{K \rightarrow \infty} \Pr(\mathcal{E}_{0,k}) \geq \lim_{K \rightarrow \infty} \left(K \left(1 - \frac{b \ln K}{K}\right)^{\frac{K}{b}} - \frac{K^2}{2} \left(1 - \frac{2}{K}\right)^{K \ln K} \right) = \lim_{K \rightarrow \infty} \left(K e^{-\ln K} - \frac{K^2}{2} e^{-2 \ln K} \right) = \frac{1}{2}. \quad (32)$$

■

C. Proof of Proposition 6

Proof: For the Rayleigh distribution we have $F_X(x) = 1 - e^{-\frac{x^2}{2\sigma^2}}$, so we simply have $\frac{1-F_X(x)}{\alpha x^\beta e^{-\lambda x^\gamma}} = 1$ for $\alpha = 1, \beta = 0, \gamma = 2$ and $\lambda = \frac{1}{2\sigma^2}$. Now observe that From l'Hôpital's [39, Page 109] role we obtain

$$\lim_{x \rightarrow \infty} \frac{1 - F_X(x)}{\alpha x^\beta e^{-\lambda x^\gamma}} = \lim_{x \rightarrow \infty} \frac{f_X(x)}{\left(\alpha \lambda \gamma - \frac{\alpha \beta}{x^\gamma}\right) x^{\beta-1+\gamma} e^{-\lambda x^\gamma}}. \quad (33)$$

For an m -Nakagami distribution with parameter $\Omega > 0$, the PDF is $f_X(x) = \frac{2m^m}{\Gamma(m)\Omega^m} x^{2m-1} e^{-\frac{m}{\Omega}x^2}$.

We choose $\alpha = \frac{m^{m-1}}{\Gamma(m)\Omega^{m-1}}, \beta = 2m - 2, \gamma = 2$ and $\lambda = \frac{m}{\Omega}$. Substituting this in (33) yields

$$\lim_{x \rightarrow \infty} \frac{f_X(x)}{\left(\alpha \lambda \gamma - \frac{\alpha \beta}{x^\gamma}\right) x^{\beta-1+\gamma} e^{-\lambda x^\gamma}} = \lim_{x \rightarrow \infty} \frac{\frac{2m^m}{\Gamma(m)\Omega^m} x^{2m-1} e^{-\frac{m}{\Omega}x^2}}{\left(\frac{2m^m}{\Gamma(m)\Omega^m} - \frac{\alpha \beta}{x^2}\right) x^{2m-1} e^{-\frac{m}{\Omega}x^2}} = 1. \quad (34)$$

For the Rice distribution we have $f_X(x) = \frac{x}{\sigma^2} e^{-\frac{x^2+v^2}{2\sigma^2}} I_0\left(\frac{xv}{\sigma^2}\right)$. We choose $\alpha = \sigma \sqrt{\frac{v}{2\pi}}, \beta = -\frac{1}{2}, \gamma = 2$ and $\lambda = \frac{1}{2\sigma^2}$. Substituting this in (33) yields

$$\lim_{x \rightarrow \infty} \frac{\frac{x}{\sigma^2} e^{-\frac{x^2+v^2}{2\sigma^2}} I_0\left(\frac{xv}{\sigma^2}\right)}{\left(\alpha \lambda \gamma - \frac{\alpha \beta}{x^\gamma}\right) x^{\beta-1+\gamma} e^{-\lambda x^\gamma}} \stackrel{(a)}{=} \lim_{x \rightarrow \infty} \frac{\frac{1}{\sigma} \sqrt{\frac{v}{2\pi}} \sqrt{x} e^{-\frac{(x-v)^2}{2\sigma^2}}}{\left(\frac{1}{\sigma} \sqrt{\frac{v}{2\pi}} + \frac{\alpha}{2x^2}\right) \sqrt{x} e^{-\frac{x^2}{2\sigma^2}}} = 1. \quad (35)$$

where (a) follows since $\lim_{x \rightarrow \infty} \frac{I_0\left(\frac{xv}{\sigma^2}\right)}{\frac{\sigma}{\sqrt{\pi x v}} e^{-\frac{xv}{2\sigma^2}}} = 1$. For a Gaussian distribution we have $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}x^2}$. We choose $\alpha = \sqrt{\frac{\sigma^2}{2\pi}}, \beta = -1, \gamma = 2$ and $\lambda = \frac{1}{2\sigma^2}$. Substituting this in (33) yields

$$\lim_{x \rightarrow \infty} \frac{f_X(x)}{\left(\alpha \lambda \gamma - \frac{\alpha \beta}{x^\gamma}\right) x^{\beta-1+\gamma} e^{-\lambda x^\gamma}} = \lim_{x \rightarrow \infty} \frac{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}x^2}}{\left(\frac{1}{\sqrt{2\pi\sigma^2}} + \frac{\alpha}{x^2}\right) e^{-\frac{1}{2\sigma^2}x^2}} = 1. \quad (36)$$

■

REFERENCES

- [1] I. Bistriz and A. Leshem, "Efficient and asymptotically optimal resource block allocation," presented at IEEE WCNC 2018.
- [2] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *Signal Processing Magazine, IEEE*, vol. 24, no. 3, pp. 79–89, 2007.
- [3] M. Hasan, E. Hossain, and D. Niyato, "Random access for machine-to-machine communication in LTE-advanced networks: issues and approaches," *IEEE Communications Magazine*, vol. 51, no. 6, pp. 86–93, 2013.
- [4] K. Akkarajitsakul, E. Hossain, D. Niyato, and D. I. Kim, "Game theoretic approaches for multiple access in wireless networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 13, no. 3, pp. 372–395, 2011.

- [5] K. Seong, M. Mohseni, and J. M. Cioffi, "Optimal resource allocation for OFDMA downlink systems," in *2006 IEEE International Symposium on Information Theory*, 2006.
- [6] W. Zhao, S. Wang, and J. Guo, "Efficient resource allocation for OFDMA-based device-to-device communication underlying cellular networks," in *Communications in China (ICCC), 2015 IEEE/CIC International Conference on*, 2015.
- [7] A. Ghosh, J. Zhang, J. G. Andrews, and R. Muhamed, *Fundamentals of LTE*. Pearson Education, 2010.
- [8] T. Yoo and A. Goldsmith, "Optimality of zero-forcing beamforming with multiuser diversity," in *Communications, 2005. ICC 2005. 2005 IEEE International Conference on*, 2005.
- [9] J.-C. Guey and L. D. Larsson, "Modeling and evaluation of MIMO systems exploiting channel reciprocity in TDD mode," in *Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th*, vol. 6, 2004, pp. 4265–4269.
- [10] D. Gesbert and M.-S. Alouini, "How much feedback is multi-user diversity really worth?" in *Communications, 2004 IEEE International Conference on*, vol. 1, 2004, pp. 234–238.
- [11] S. Sanayei and A. Nosratinia, "Opportunistic downlink transmission with limited feedback," *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 4363–4372, 2007.
- [12] J. Chen, R. A. Berry, and M. L. Honig, "Large system performance of downlink OFDMA with limited feedback," in *Information Theory, 2006 IEEE International Symposium on*, 2006, pp. 1399–1403.
- [13] —, "Limited feedback schemes for downlink OFDMA based on sub-channel groups," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1451–1461, 2008.
- [14] J. Leinonen, J. Hämäläinen, and M. Juntti, "Performance analysis of downlink OFDMA resource allocation with limited feedback," *IEEE Transactions on Wireless Communications*, vol. 8, no. 6, pp. 2927–2937, 2009.
- [15] P. Erdos and A. Renyi, "On random matrices," *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, vol. 8, no. 455-461, p. 1964, 1964.
- [16] E. Yaacoub and Z. Dawy, "A survey on uplink resource allocation in OFDMA wireless networks," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 2, pp. 322–337, 2012.
- [17] S. Sadr, A. Anpalagan, and K. Raahemifar, "Radio resource allocation algorithms for the downlink of multiuser OFDM communication systems," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 3, pp. 92–106, 2009.
- [18] L. M. Hoo, B. Halder, J. Tellado, and J. M. Cioffi, "Multiuser transmit optimization for multicarrier broadcast channels: asymptotic FDMA capacity region and algorithms," *IEEE Transactions on Communications*, vol. 52, no. 6, pp. 922–930, 2004.
- [19] K. Kim, Y. Han, and S.-L. Kim, "Joint subcarrier and power allocation in uplink OFDMA systems," *IEEE Communications Letters*, vol. 9, no. 6, pp. 526–528, 2005.
- [20] W. Yu and J. M. Cioffi, "Constant-power waterfilling: performance bound and low-complexity implementation," *IEEE Transactions on Communications*, vol. 54, no. 1, pp. 23–28, 2006.
- [21] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *Selected Areas in Communications, IEEE Journal on*, vol. 20, no. 5, pp. 1105–1115., 2002.
- [22] I. C. Wong and B. L. Evans, "Optimal downlink OFDMA resource allocation with linear complexity to maximize ergodic rates," *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, pp. 962–971, 2008.
- [23] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE Journal on selected areas in communications*, vol. 21, no. 2, pp. 171–178, 2003.

- [24] D. J. Love, R. W. Heath, V. K. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE Journal on selected areas in Communications*, vol. 26, no. 8, pp. 1341–1365, 2008.
- [25] Y. S. Al-Harathi, A. H. Tewfik, and M.-S. Alouini, "Multiuser diversity with quantized feedback," *IEEE Transactions on Wireless Communications*, vol. 6, no. 1, pp. 330–337, 2007.
- [26] I. Toufik and H. Kim, "MIMO-OFDMA opportunistic beamforming with partial channel state information," in *Communications, 2006. ICC'06. IEEE International Conference on*, vol. 12, 2006, pp. 5389–5394.
- [27] P. Svedman, S. K. Wilson, L. J. Cimini, and B. Ottersten, "Opportunistic beamforming and scheduling for OFDMA systems," *IEEE Transactions on Communications*, vol. 55, no. 5, pp. 941–952, 2007.
- [28] O. Naparstek and A. Leshem, "A fast matching algorithm for asymptotically optimal distributed channel assignment," in *Digital Signal Processing (DSP), 2013 18th International Conference on*, 2013.
- [29] —, "Fully distributed optimal channel assignment for open spectrum access," *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 283–294, 2014.
- [30] —, "Expected time complexity of the auction algorithm and the push relabel algorithm for maximum bipartite matching on random graphs," in *Random Structures & Algorithms*. Wiley Online Library, 2014.
- [31] I. Bistriz and A. Leshem, "Asymptotically optimal distributed channel allocation: a competitive game-theoretic approach," in *Communication, Control, and Computing (Allerton), 2015 53rd Annual Allerton Conference on*, 2015.
- [32] C. H. Papadimitriou and K. Steiglitz, *Combinatorial optimization: algorithms and complexity*. Courier Corporation, 1998.
- [33] D. H. Lehmer, *Applied Combinatorial Mathematics*, E. F. Beckenbach, Ed. John Wiley & Sons, 1964.
- [34] A. B. Siddique, S. Farid, and M. Tahir, "Proof of bijection for combinatorial number system," arXiv preprint arXiv:1601.05794. 2016.
- [35] I. Bistriz and A. Leshem, "Game theoretic dynamic channel allocation for frequency-selective interference channels," arXiv:1705.02957.
- [36] E. LTE, "Evolved universal terrestrial radio access (e-utra); base station (bs) radio transmission and reception (3gpp ts 36.104 version 8.6. 0 release 8), july 2009," *ETSI TS*, vol. 136, no. 104, p. V8, 2009.
- [37] I. Bistriz and A. Leshem, "Game theoretic resource allocation for m-dependent channel with application to OFDMA," in *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*, 2017, pp. 4316–4320.
- [38] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
- [39] W. Rudin *et al.*, *Principles of mathematical analysis*. McGraw-hill New York, 1964.